

(12)

**EUROPEAN PATENT APPLICATION**

(21) Application number: **87402542.2**

(51) Int. Cl.<sup>4</sup>: **C 12 N 15/00**

(22) Date of filing: **10.11.87**

(30) Priority: **11.11.86 JP 268362/86**

(43) Date of publication of application:  
**18.05.88 Bulletin 88/20**

(84) Designated Contracting States: **DE FR GB IT**

(71) Applicant: **MITSUBISHI CHEMICAL INDUSTRIES LIMITED**  
**5-2, Marunouchi 2-chome Chiyoda-ku**  
**Tokyo 100 (JP)**

(72) Inventor: **Shibui, Tatsuro**  
**M606-11 San-Howaito 6-15-4 Minami-Naruse**  
**Machida-shi Tokyo (JP)**

**Kamizono, Michiru**  
**F202-24 San-Howaito 2-8-11 Tsuruma**  
**Machida-shi Tokyo (JP)**

**Teranishi, Yutaka**  
**1-14-2 Takane**  
**Sagamihara-shi Kanagawa-ken (JP)**

(74) Representative: **Gutmann, Ernest et al**  
**S.C. Ernest Gutmann - Yves Plasseraud 67, boulevard**  
**Haussmann**  
**F-75008 Paris (FR)**

(54) **Hybrid promoter, expression controlling DNA sequence and expression vector.**

(57) The present invention provides novel hybrid promoters comprising an RNA polymerase recognition region having a base sequence equivalent to that of the RNA polymerase recognition region of a phage promoter, and an RNA polymerase binding region having a base sequence equivalent to that of the RNA polymerase binding region of an *E. coli* promoter. Disclosed also are expression-controlling DNA sequences comprising said hybrid promoter and an operator region having a base sequence equivalent to that of the operator region of *E. coli*, which allow for efficient expression in a well-controlled manner. Disclosed further are expression vectors comprising a terminator and the gene encoding a target protein as well as said expression-controlling DNA sequence.

**EP 0 267 851 A2**

## Description

## HYBRID PROMOTER, EXPRESSION CONTROLLING DNA SEQUENCE AND EXPRESSION VECTOR

The present invention relates to a hybrid promoter which allows for efficient expression, to an expression-controlling DNA sequence comprising said hybrid promoter and to an expression vector comprising said expression-controlling DNA sequence.

Expression of genes takes place in two steps, i.e., transcription and translation. In transcription step an RNA polymerase specifically recognizes "-35 region" located approximately 35 base pairs upstream of the start of mRNA synthesis within a region, called promoter, which spans about 50 base pairs and is located upstream of a structural gene, and then binds to "-10 region" located about 10 base pairs upstream of the start of mRNA synthesis so as to start mRNA synthesis which is stopped at a site called terminator. The subsequent translation step is a protein synthesizing process according to triplet codons, which is initiated when a ribosome and the like recognize and bind to both a specific sequence called SD (Shine-Dalgarno) sequence and an initiation codon such as AUG on mRNA and is stopped at a termination codon.

Except for phenomena peculiar to higher organisms, such as splicing, the major part of the species specificities in gene expression exists in the specific sequences at the starting points of both transcription and translation and in the recognition processes thereof, whereas the subsequent elongation reactions generally have no specificity. Accordingly, a desired or target protein can be synthesized by incorporating a region including the promoter and the initiation codon which are responsible for the initiation of gene expression, into a plasmid cloning vector and then ligating the structural gene encoding the desired protein downstream to the region.

On expressing the desired protein using such an expression vector as described above, expression efficiency is very critical, especially from an industrial viewpoint.

There have hitherto been proposed various expression vectors using promoters such as a *tac* promoter, *P<sub>LP</sub>* promoters and an expression-controlling DNA sequence consisting of a T5 phage promoter (-35 region and -10 region) and a *lac* operator (Japanese Patent Application Laying Open No. 61-181386).

The present invention provides an expression vector which enables efficient expression in a well-controlled manner. The present invention also provides a hybrid promoter which has, as an RNA polymerase recognition region, a base sequence equivalent to the base sequence of a promoter derived from a phage and, as an RNA polymerase binding region, a base sequence equivalent to the base sequence of a promoter derived from *E. coli*. The present invention further provides an expression-controlling DNA sequence comprising said hybrid promoter and an operator region, a base sequence of which is equivalent to the base sequence of an operator region derived from *E. coli*.

Figure 1 shows schematically the method of construction of an expression vector according to the present invention.

The hybrid promoter of the present invention comprises preferably the -35 region which the RNA polymerase recognizes and the -10 region to which it is bound. The -35 region has the base sequence equivalent to that of the -35 region of the phage promoter. The -10 region has the base sequence equivalent to that of the -10 region of the *E. coli* promoter. By "an equivalent base sequence" is meant in this specification either a base sequence that is entirely the same as the naturally occurring one or any of the base sequences that are partially modified by substitution, insertion or deletion to such an extent that its original functions are still retained. As for the -35 region of the present invention, 5 - 12 base pair fragments comprising at least the base pairs shown by (a) below may be preferably used.



Among said fragments, there may be mentioned the base sequence corresponding to the -35 regions in the promoters of P25, P26, P28 or P207 of T5 phage.

As the -10 region of the present invention, preferably 7 - 13 base pair fragments comprising at least the base pairs shown by (b) below, more preferably the fragment shown by (c) below may be used.

-10

5' T A T A A T G 3' ..... (b)  
3' A T A T T A C 5'

-10

5' T A T A A T G T G T G G 3' ..... (c)  
3' A T A T T A C A C A C C 5'

The base sequence corresponding to the -10 region in the promoters of lac UV5, rrnD<sub>1</sub> or rrnE<sub>1</sub> of E. coli may be exemplified.

According to the present invention, the -35 region is located upstream of the -10 region, wherein both regions may preferably be bound to each other through any optional base sequence of about 10 - 20 base pairs although it is also possible to have them ligated directly with each other.

The expression-controlling DNA sequence of the present invention comprises the above-mentioned hybrid promoter and the base sequence equivalent to that of the operator region derived from E. coli.

The base sequence of corresponding to the regions of lac UV5, trp or rec A of E. coli may be illustrated.

5' AATTGTGAGCGGATAACAATT 3' ..... (d)

3' TTAACACTCGCCTATTGTAA 5'

For example, any DNA fragment comprising the base pairs shown by (d) below may be used as an operator region.

The operator region of the present invention is normally incorporated in the neighborhood of the -35 region and the -10 region, though its location is by no means restricted thereto. It is preferably located downstream of the -10 region via any optional base sequence of 0 - 10 base pairs and is more preferably located immediately downstream of the -10 region.

The above-mentioned -35 region, and -10 region and operator region may be respectively prepared from the above phage, and E. coli, or alternatively may be prepared by chemical synthesis.

The expression-controlling DNA sequence is normally used in combination with, for example, SD sequence which is a ribosome binding region, an initiation codon, a termination codon and a repressor region. It is especially preferable to combine immediately downstream from the initiation codon a (multi-) cloning site in 6 - 50 nucleotide length consisting of a DNA sequence providing at least one or, preferably, a plurality of restriction enzyme sites which may be advantageously used in insertion of the genes encoding various target proteins. Interferons of every kind, IL-2, apolipoproteins A and E, TNF, cardiodilatin, cardionatrin and the like may be illustrated as unlimited examples of the target protein.

The above-mentioned (multi-) cloning site is illustrated, for example, by a DNA fragment having a plurality of restriction sites as shown below.

Kpn I      Sma I      Apa I      Sal I      Bgl II      Ava I

5' ATGGTACCCGGGCCCCGTCGACAGATCTGGG-  
3' TACCATGGGCCCCGGGCAGCTGTCTAGACCC-

← fMet

GCA 3'  
CGTCTAG 5'

The expression vector of the present invention comprises the gene encoding the target protein and a terminator inserted downstream of the expression-controlling DNA sequence described above. Among the terminators there may be mentioned the 3' region of lipoprotein (lpp) gene, the terminators of the trp operon and rRNA gene, and the t<sub>0</sub> of λ phage.

The hybrid promoter and the expression-controlling DNA sequence of the present invention make it possible to efficiently synthesize the target protein in a well-controlled fashion.

The present invention will be described more in detail, but without limitation, in the following example.

Example:

## (1) Construction of an expression-controlling DNA sequence

A DNA fragment consisting of the same sequences as those of the -10 region, the operator region and the transcription initiation point of the *lac* UV5 promoter, and a DNA fragment consisting of the same sequence as that of the -35 region of the T5 phage P25 promoter ((1) - (4) shown below) were synthesized using a DNA synthesizer Model 308 A (Nikkaki Co., Japan)

- ① 5' AATTCAAAAATTTATTTGCTTTCAGGAAAAATTTTCT3'  
 ② 5' ATACAGAAAAATTTTCCTGAAAGCAAATAAATTTTGT3'  
 ③ 5' GTATAATGTGTGGAATTGTGAGCGGATAACAATTCGGATCCA3'  
 ④ 5' AGCTTGGATCCGAAATTGTTATCCGCTCACAAATCCACACATT3'

E c o R

-35

① 5' AATTCAAAAATTTATTTGCTTTCAGGAAAAATTTTCT3'  
 3' GTTTTTAAATAAACGAAAGTCCTTTTAAAAAGACATA5' ②

-10                      Operator                      BamHI HindIII

③ 5' GTATAATGTGTGGAATTGTGAGCGGATAACAATTCGGATCCA3'  
 3' TTACACACCTTAACACTCGCCTATTGTTAAAGCCTAGGTTCCA5' ④

(1) 37 bases: a DNA fragment corresponding to the DNA fragment comprising the -35 region of T5 phage P25 promoter.

(2) 37 bases: a DNA fragment complementary to (1).

(3) 43 bases: a DNA fragment corresponding to the DNA fragment comprising -10 region and the operator region of *lac* UV5.

(4) 43 bases: a DNA fragment complementary to (3).

Each synthesized DNA fragment in an amount of 100 µg was incubated in 2 ml of 28% ammonium aqueous solution at 55°C for 6 hours for deprotection and purified by 17% acrylamide - 7M urea gel electrophoresis so as to obtain 10 µg of each purified DNA fragment. Subsequently, 1 µg of each DNA fragment was incubated with one unit of T4 polynucleotidekinase in 10 µl of a buffer solution containing 10mM Tris-HCl, pH 7.6 and 10mM MgCl<sub>2</sub> at 37°C for one hour in order to phosphorylate the 5'-terminus thereof. After addition of NaCl to the buffer solution to a final concentration of 50mM, the DNA fragment (1) was mixed with the fragment (2), and the fragment (3) was mixed with the fragment (4), respectively. Then, each mixture was heated at 100°C for 3 minutes and gradually cooled for annealing to result in 2 µg each of double-stranded DNA fragments of 37 bp and 43 bp, respectively.

The double-stranded DNA fragments of 0.5µg each were inserted together into 1 µg of plasmid pBR 322 treated with *Eco* RI and *Hind* III using 1 unit of T4 DNA ligase in 10 µl of a buffer solution containing 10mM Tris-HCl, pH 7.5, 6mM MgCl<sub>2</sub>, 10mM ATP and 1mM DTT at 4°C for 16 hours.

To this reaction mixture were added CaCl<sub>2</sub> to a final concentration of 100mM. The resultant mixture was mixed with competent cells of *E. coli* HB101 and incubated at 0°C for 10 minutes and then at 37°C for 5 minutes to carry out transformation of *E. coli* HB101. The transformant thus obtained acquired tetracycline resistance and could be easily selected by adding tetracycline in a final concentration of 20 µg/ml to a medium.

Plasmid pBRpac with *pac* promoter activity was thus obtained. The base sequence of the *pac* promoter was determined by a conventional method, i.e., the dideoxy method, which is referred to in detail in Anal. Biochem Vol. 152,232. The competent cells may be obtained by collecting, in a conventional manner, the cells in their logarithmic phase, suspending them in a 1/3 volume of 0.1M CaCl<sub>2</sub> at 0°C, allowing the suspension to stand for 30 minutes, and recovering the cells in the suspension and again suspending them in a 1/100 volume of 0.1M CaCl<sub>2</sub>.

## (2) Incorporation of a terminator into the expression-controlling DNA sequence

One µg of the plasmid pBRpac (4.4 kbp) obtained above was digested with 1 unit each of *Bam*HI and *Pvu*II in 10 µl of a buffer solution containing 10mM Tris-HCl, pH 7.5, 100mM NaCl and 6mM MgCl<sub>2</sub> by incubating the mixture at 37°C for 2 hours. The reactant was precipitated with ethanol and evaporated to dryness for preservation.

On the other hand, 5 µg of plasmid pINIII A1 (The EMBOJ, Vol. 3, 10, 2437, 1984) was digested with 5 units *Kpn*I in 50 µl of a buffer solution containing 10 mM Tris-HCl, pH 7.5 and 6mM MgCl<sub>2</sub> by incubating the mixture at 37°C for 2 hours. To the mixture was added 5.5 µl of a buffer solution containing 330 mM Tris-Ac, pH7.9, 660mM KAc and 100mM Mg (Ac)<sub>2</sub> and the protruding 3' termini of the digested pINIII A1 were removed by incubation with 10 units of T4DNA polymerase at 37°C for 15 minutes. The polymerase was inactivated by heating at 70°C for 10 minutes. To this reaction mixture was then added 6.0 µl of a buffer solution containing 100mM Tris-HCl, pH 7.5, 1M NaCl and 60mM MgCl<sub>2</sub> and digestion with 5 units of *Bam*HI was carried out at 37°C

for 2 hours, followed by the addition of 60  $\mu$ l of an aqueous saturated phenol solution for extraction of proteins. After removal of proteins, ethanol precipitation was repeated three times to result in 0.5  $\mu$ g of *lpp* 3' region (~500 bp) fragment comprising the terminator. This DNA fragment was inserted into said digested plasmid (pBRpac) using one unit of T4DNA ligase in 10  $\mu$ l of a buffer solution containing 10mM Tris-HCl, pH7.5, 10mM MgCl<sub>2</sub>, 10mM ATP and 1mM DTT by incubating the mixture at 4°C for 16 hours, resulting in plasmid pPac.

The plasmid pPac may be stably maintained in JM 105 or JM 109 which is *lac* I<sup>-</sup> strain. Thus 1  $\mu$ g of pPac was used to transform JM 105 strain according to the method described in Example, (1). The resulting transformants were grown in 1 litre of a minimal medium containing 20 mg/ml of ampicillin and 1 mg of pPac was purified in a conventional manner.

### (3) Incorporation of a multi-linker

A multi-linker, i.e., multi-cloning site, combined with SD sequence and the initiation Met sequence, as shown in Fig. 1, was synthesized using Model 308A (Nikkaki Co.) and purified according to the method described in Example, (1) to obtain 10  $\mu$ g of DNA.

On the other hand, 1  $\mu$ g of the aforementioned plasmid pPac was cleaved with 1 unit of BamHI in 10  $\mu$ l of a buffer solution containing 10mM Tris-HCl, pH7.5, 100 mM NaCl and 6mM MgCl<sub>2</sub> at 37°C for 2 hours, precipitated with ethanol and evaporated to dryness for preservation.

This cleaved plasmid pPac and 1  $\mu$ g of the above multi-linker were incubated with 1 unit T4DNA ligase in 10  $\mu$ l of a buffer solution containing 10mM Tris-HCl, pH 7.5, 10mM MgCl<sub>2</sub>, 10mM ATP and 1mM DTT at 4°C for 16 hours to obtain plasmid pMT2 with the multi-linker inserted therein.

### (4) Incorporation of a *lac* repressor gene

10  $\mu$ g of plasmid pMC9 (Proc. Natl. Acad. Sci. USA, 80, 3015 (1983)) was cleaved with 10 units of EcoRI in 100  $\mu$ l of a buffer solution containing 100mM Tris-HCl, pH 7.5, 10mM NaCl and 6mM MgCl<sub>2</sub> by incubating the mixture at 37°C for 2 hours, precipitated with ethanol and then evaporated to dryness for preservation. As a result, 2.5  $\mu$ g of 1.7 Kbp DNA fragment comprising *lac* I, i.e., a repressor gene was obtained.

On the other hand, 1  $\mu$ g of the aforementioned plasmid pMT2 was cleaved with 1 unit EcoRI in 10  $\mu$ l of a buffer solution containing 100mM Tris-HCl, pH 7.5, 10mM NaCl and 6mM MgCl<sub>2</sub> by incubating the mixture at 37°C for 2 hours.

The cleaved plasmid pMT2 and 1  $\mu$ g of the 1.7 Kbp DNA fragment were ligated with each other by the use of 1 unit of T4DNA ligase in 20  $\mu$ l of a buffer solution containing 10mM Tris-HCl pH 7.5, 10mM MgCl<sub>2</sub>, 10mM ATP and 1mM DTT to give plasmid pMTI2.

### (5) Construction of an expression plasmid

After cleaving 1  $\mu$ g of pMTI2 in 10  $\mu$ l of a buffer solution containing 10mM Tris-HCl pH7.5, 6mM MgCl<sub>2</sub> and 1 unit Kpn I at 37°C for 2 hours, 1  $\mu$ l of a buffer solution containing 330mM Tris-Ac pH 7.9, 660mM KAc and 100mM Mg (Ac)<sub>2</sub> was added to the reaction mixture. The protruding 3' termini of the digested pMTI2 were then removed by treatment with 1 unit of T4DNA polymerase in the reaction mixture at 37°C for 15 minutes and the polymerase was inactivated by heating at 70°C for 10 minutes. To the reaction mixture was further added 1.2  $\mu$ l of a buffer system containing 100mM Tris-HCl, pH7.5, 1M NaCl and 60mM MgCl<sub>2</sub> and digestion with 1 unit of Bgl II was carried out at 37°C for 2 hours. After treatment with phenol for removal of proteins, ethanol precipitation was repeated three times to result in precipitate, which was then evaporated to dryness for preservation.

Subsequently, 1  $\mu$ g of the HpaI-BglII fragment (219 bp) of the gene of AVD (Arterial Vaso Dilatin), i.e., an atriavasodilation peptide, and the plasmid pMTI2 treated as above, were linked with each other in 10  $\mu$ l of a buffer solution containing 10mM Tris-HCl, pH7.5, 10mM MgCl<sub>2</sub>, 10mM ATP and 1mM DTT using 1 unit of T4DNA ligase to obtain plasmid pMTAVD. *E. coli* YA21 was transformed with 0.1  $\mu$ g of this pMTAVD according to the method described in Example, (1).

The resulting transformant was capable of producing approximately 36 mg/l of AVD within the cells as a result of the expression of AVD, induced by adding IPTG (isopropyl-1-thio- $\beta$ -D-galactoside).

The above mentioned AVD gene was obtained by digesting with HpaI and BglII plasmid pANFSHB which had been acquired from pANF66 by such a site-specific mutation that a HpaI site and a BglII site were introduced 5' and 3' to the AVD gene, respectively (Nature, 310, 699, (1985)).

## Claims

1. A hybrid promoter comprising an RNA polymerase recognition region having a base sequence equivalent to that of the RNA polymerase recognition region of a phage promoter, and an RNA polymerase binding region having a base sequence equivalent to that of the RNA polymerase binding region of an *E. coli* promoter.

2. A hybrid promoter as claimed in claim 1 wherein the RNA polymerase recognition region comprises -35 region.

3. A hybrid promoter as claimed in claim 2 wherein the -35 region comprises at least the following base

sequence:

TTGCT  
AACGA .

5 4. A hybrid promoter as claimed in claim 1 wherein the RNA polymerase recognition region has a base sequence equivalent to that of the RNA polymerase recognition region of a T5 phage P25 promoter.

5. A hybrid promoter as claimed in claim 1 wherein the RNA polymerase binding region comprises -10 region.

6. A hybrid promoter as claimed in claim 5 wherein the -10 region comprises at least the following base sequence:

10 TATAATG  
ATATTAC.

7. A hybrid promoter as claimed in claim 6 wherein the -10 region comprises at least the following base sequence:

15 TATAATGTGTGG  
ATATTACACACC .

8. A hybrid promoter as claimed in claim 1 wherein the RNA polymerase binding region has a base sequence equivalent to that of the RNA polymerase binding region of a *lac* UV5 promoter.

9. An expression-controlling DNA sequence comprising the hybrid promoter according to claim 1, and an operator region having a base sequence equivalent to that of an *E. coli* operator region.

20 10. An expression-controlling DNA sequence as claimed in claim 9 wherein the operator region has a base sequence equivalent to that of a *lac* UV5 operator region.

11. An expression-controlling DNA sequence as claimed in claim 9 comprising a DNA sequence having an initiation codon and at least one restriction enzyme site located immediately downstream of said initiation codon.

25 12. An expression-controlling DNA sequence as claimed in claim 11 comprising the following base sequence:

AAATCAAAAATTTATTTGCTTTCAGGAAAATTTTCTGT  
GTTTTAAATAAACGAAAGTCCTTTTAAAAAGACA

30 ATAATGTGTGGAATTGTGAGCGGATAACAATTCGGATCCA  
TATTACACACCTTAACACTCGCCTATTGTTAAAGCCTAGGTTCGA .

13. An expression vector wherein a gene encoding a target protein and a terminator are inserted downstream of the expression-controlling DNA sequence according to claim 9.

35 14. An expression vector as claimed in claim 13 wherein the operator region has a base sequence equivalent to that of a *lac* UV5 operator region.

15. An expression vector as claimed in claim 13 comprising a DNA sequence having an initiation codon and at least one restriction enzyme site located immediately downstream of said initiation codon.

40

45

50

55

60

65

**Fig.1**  
**(i)**

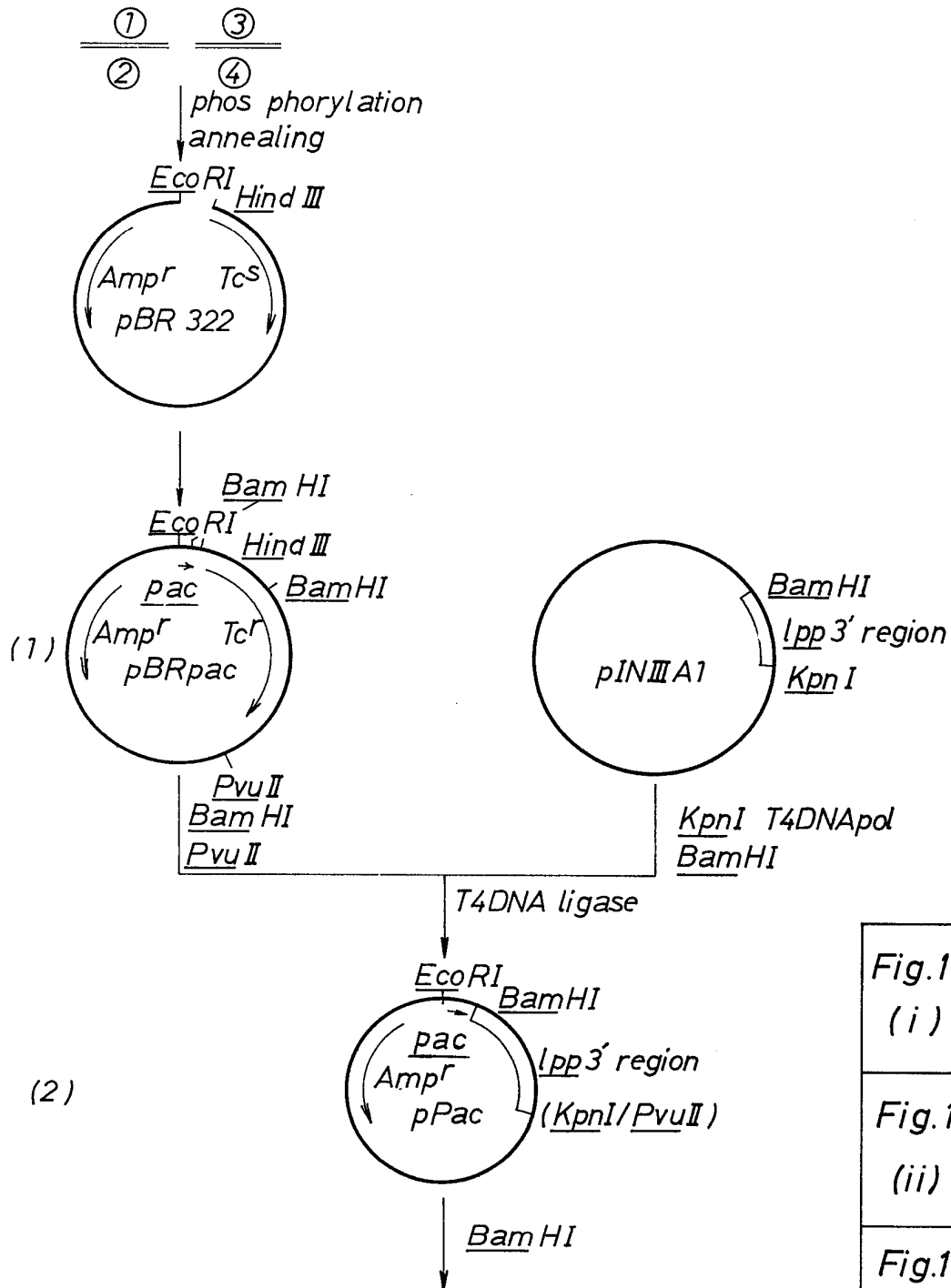


Fig.1  
(i)

Fig.1  
(ii)

Fig.1  
(iii)

**Fig.1**  
**(ii)**

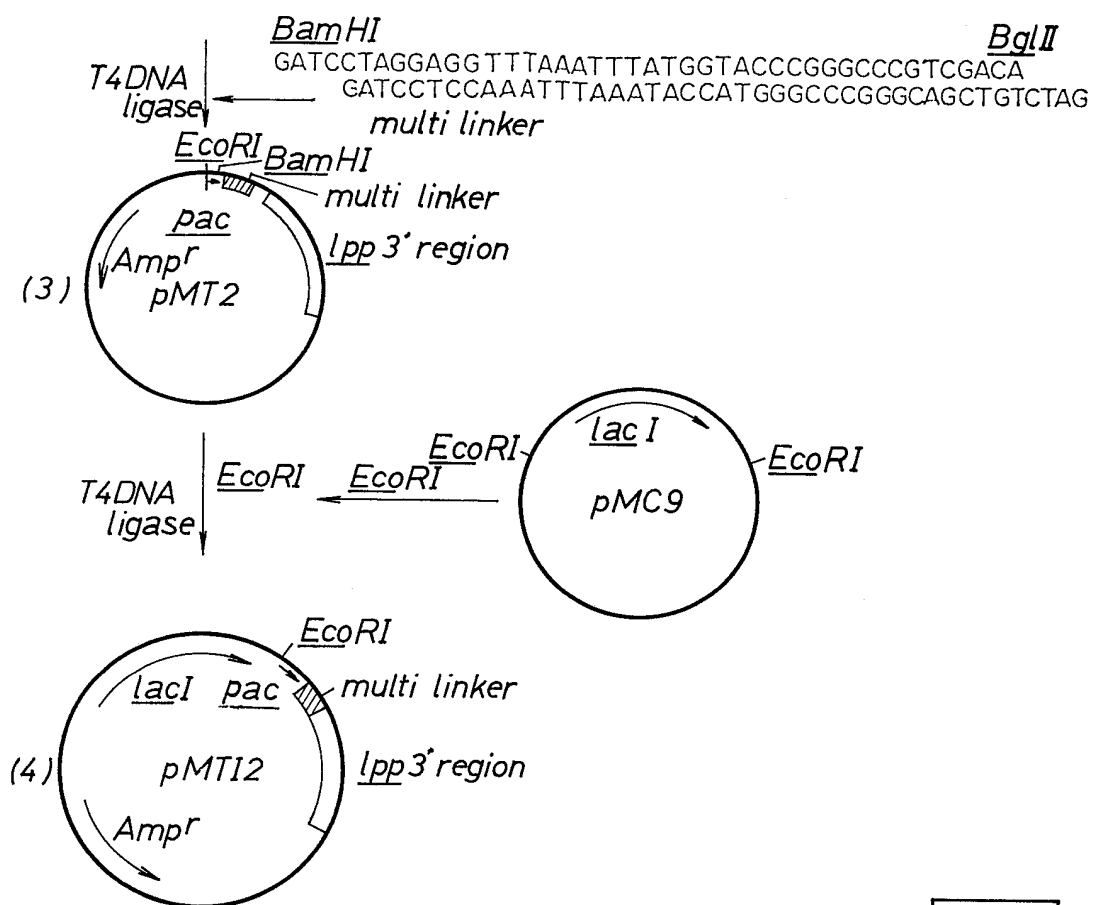


Fig.1  
(i)

Fig.1  
(ii)

Fig.1  
(iii)



# Fig.1 (iii)

0267851

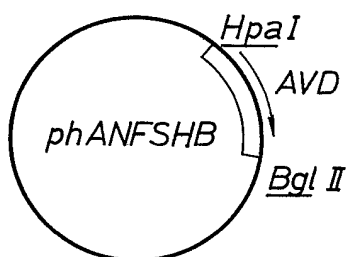
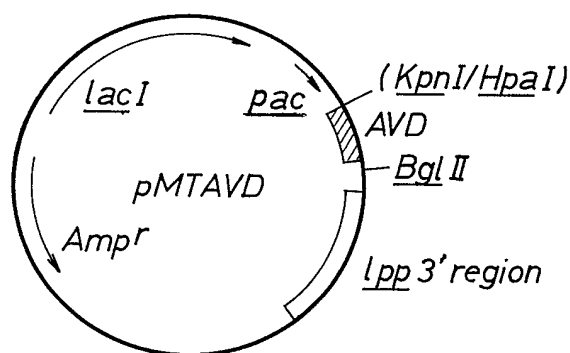
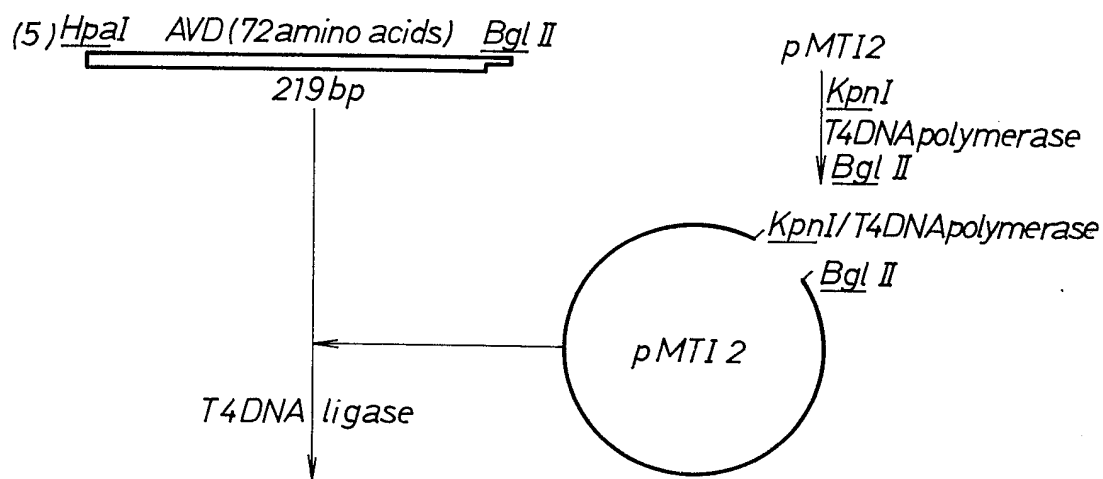


Fig.1  
(i)

Fig.1  
(ii)

Fig.1  
(iii)